

# Inference in Hybrid Bayesian Networks

Carlos Badenes

Computational Intelligence Group, Departamento de Inteligencia Artificial, Universidad Politécnica de Madrid, Spain

## Abstract

There are many interesting domains containing not only discrete variables, but also continuous values such as distance, temperature or location. Hybrid models are used for representing uncertainty in these type of domains but it is well-known that only for linear Gaussian models an exact inference is possible. Even so, the complexity of these algorithms is so high that present significant challenges to perform. In this article, we review the advances that have been development in this regard in the last years.

KEY WORDS: Hybrid Bayesian Networks; Inference;

## 1 Introduction

The introduction of continuous variables in a graphical model has several particularities. The representation of factors that imply continuous variables is the first challenge. It is impossible to represent a factor over continuous variables in a general way, so you have to select a valid representation for each Conditional Probability Distribution (CPD) in the network. It is generally unlikely that you can find a single parametric family that can correctly encode all of the intermediate factors in the network. (Koller and Friedman (2009))

The second challenge is using integration rather than summation when you marginalize a variable because not all functions are integrable: in some cases, the integral may be infinite or even ill defined. In addition, even functions where the integral is well defined may not have a closed-form integral, requiring the use of a numerical integration method, which is usually approximate.

Thus, inference in hybrid models, although similar to discrete inference, implies a new set of challenges. When the distribution is a multivariate Gaussian many of these challenges disappear. The integration operation used to marginalization is always well defined, and it is guaranteed to produce a finite integral under certain conditions.

The remainder of this article is organized as follows. Section State-of-the-art describes articles, papers and researches centered in how to inference in Hybrid Bayesian Networks (HBN). After that, in the section Conclusions and future research, the main open lines about this are mentioned.

## 2 State-of-the-art

Since Bayesian network (BN) was introduced in the field of artificial intelligence in 1980s, some inference algorithms have been developed for probabilistic reasoning. However, when continuous variables are present in this networks, their dependence relationships could be nonlinear and their probability distributions could be arbitrary.

Heskes and Zoeter (2003) applied a generalized belief propagation to approximate inference in HBN. They replaced the strong marginalization in discrete networks:  $P_{\alpha}(x_{\beta}) = \sum_{X_{\alpha \setminus \beta}} P_{\alpha}(X_{\alpha})$  with a *weak* marginal in HBN. Changing summation into integration these networks consisting of continuous Gaussian variables can be handled in a similar manner. Continuous Gaussian potentials and beliefs

are summarized with a mean  $E[x]$ , covariance matrix  $E[(x - E[x])(y - E[y])]$ , and (if necessary) proportionality constant.

But Schrempf and Hanebeck (2004) considered that using Continuous Gaussian potentials is a drawback in the mere usage of the first two moments (mean and variance) to characterize continuous densities because exist different densities having identical first moments. Another problem detected is can not handle discrete nodes as children of continuous parents. An attempt to remove this restriction by using sigmod functions and it is picked up to include it into Lauritzen’s mechanism. But again, their accuracy is restricted to the first two moments of the densities. They proposed using hybrid conditional densities to capture the relationship between continuous and discrete variables. These densities must describe the probability of a continuous or discrete random variable, depending on the state of a set of mixed parent variables. Mixed means the set of parent variables contains continuous and discrete variables as well. The conditional densities  $f^*(x|u_1, \dots, u_n)$  are modeled using Gaussian mixtures in the continuous case and as sum over Gaussians and Dirac pulses in the case that  $x$  is discrete. One drawback of this approach is the limitation to singly connected graphs. One possibly way to overcome this is to find a method to do clustering in the graph.

Existing other researches focuses on special instances, such as *Conditional Linear Gaussian* (CLG) (Lauritzen (1992)) where a discrete node can have continuous children, but a continuous node is not allowed to have discrete child and all the local probability models of continuous variables are conditional linear Gaussian CPDs.

Cobb and Shenoy (2004) considered that when a HBN has variables that are a deterministic function of its parents and some of the parents are continuos, then the joint density function does not exist. *Conditional linear Gaussian* (CLG) distributions can handle such cases when the deterministic function is linear and continuous variables are normally distributed. But when continuous variables are not normally distributed then the approximate inference can be performed using *Mixtures of Truncated Exponentials* (MTE) potential.

Later, Cobb and Shenoy (2005) used MTE to approximate any probability density function (PDF) so can always be marginalized in a closed form.

For Bayesian networks with arbitrary continuous variables Shenoy (2006) introduce a method for exact inference. This method consists of approximating general HBNs by a mixture of Gaussians (MoG) BNs. Then it allows to use the Lauritzen-Jensen (LJ) algorithm for a bigger class of HBNs with continuous chance nodes with non-Gaussian distributions, networks with no restrictions on the topology of discrete and continuous variables, with conditionally deterministic variables that are a nonlinear function of their continuous parents, and with continuous chance variables whose variances are functions of their parents. This method first approximates a given HBN by a MoG BN, and then using the LJ method for exact inference in MoG BNs.

On the other hand, Sun and Chang (2007) propose an approximate inference algorithm called *Unscented Message Passing* (UMP-BN) that combines a deterministic sampling method and Pearl’s message passing algorithm to provide the estimates of the first two moments of the posterior distributions. In this approximation, messages are represented by mean and variance of the continuous distribution for every node.

But in a general HBN with nonlinear and/or non Gaussian variables there is no existing method that could produce exact posterior distribution. Yuan and Druzel (2006) propose an algorithm called *Hybrid Loopy Belief Propagation* (HLBP) which extends the *Loopy Belief Propagation* (LBP) and *Nonparametric Belief Propagation* (NBP) algorithms to deal with these general networks. The main idea is to represent the LBP messages with mixture of Gaussians and formulate their calculation as Monte Carlo integration problems.

Later, Yuan and Drudzel (2007) describe an importance sampling-based algorithm that directly deals with evidential reasoning in general HBN. They propose a technique called *delayed importance function generation* that applies the HLBP to calculate the importance function and propose another technique called *soft arc revesal* to draw samples when a deterministic variable has been observed.

This technique makes importance sampling a viable approach for hybrid models. More specifically, they extend the EPIS-BN (Yuan and Druzdzel (2006)) to the most general setting.

Sun et al. (2010) postulate in their paper a new inference approach called *Direct Message Passing for Hybrid Bayesian Network* (DMP-HBN) by unifying message passing between different types of variables. This algorithm is able to provide an exact solution for polytree networks, and approximate solution by loopy propagation for general hybrid models. Since DMP-HBN is a distributed algorithm utilizing only local information, there is no need to transform the network structure as required by the JT. In this model is not required to partition the hybrid model into different network segments, each containing only pure type of variables and then conduct message passing separately. Now, messages can be exchanged directly between discrete and continuous variables within an unified framework.

### 3 Conclusions and future research

In this short state-of-the-art we have reviewed some of the most important research papers on inference in Hybrid Bayesian Networks published to date and the algorithms proposed to revolve it. The quality of these algorithms, quality as the performance, is constantly evolving and improving. Being this aspect the main line of research actually.

### References

- S. Lauritzen, *Propagation of probabilities, means and variances in mixed graphical association models* JASA, vol. 87, no. 420, pp. 1098– 1108, 1992.
- Heskes T. and Zoeter O., *Generalized belief propagation for approximate inference in hybrid Bayesian networks* SNN, University of Nijmegen, Nijmegen, The Netherlands, 2003.
- Schrempf O. and Hanebeck U., *A New Approach for Hybrid Bayesian Networks Using Full Densities* Institute of Computer Design and Fault Tolerance, Universitat Karlsruhe, 2004.
- Cobb B. R. and Shenoy P. P., *Inference in Hybrid Bayesian Networks with Deterministic Variables* School of Business, Universitat of Kansas, 2004.
- Cobb B. R. and Shenoy P. P., *Nonlinear deterministic relationships in Bayesian networks* In Proceedings of Eighth European Conference on Symbolic and Quantitative Approaches to Reasoning with uncertainty, (ECSQARU-05), pages 27-38, Berlin, 2005
- Cobb B. R. and Shenoy P. P., *Hybrid Bayesian Networks with Linear Deterministic Variables* Department of Economics and Business, Virginia Military Institute, Lexington, 2005
- Cogate V. and Detcher R., *Approximate inference algorithms for hybrid Bayesian networks with discrete constraints* In Proceedings of 21th Annual Conference on Uncertainty in Artificial Intelligence, (UAI-05), AUAI Press Corvallis, Oregon, 2005
- Shenoy P. P., *Inference in Hybrid Bayesian Networks using Mixtures of Gaussians* University of Kansas School of Business, Lawrence, KS, 2006
- Yuan C. and Druzdzel M. J., *Importance sampling algorithms for Bayesian Networks: Principles and Performance* Mathematical and Computer Modelling, 43:1189-1207, 2006
- Yuan C. and Druzdzel M. J., *Hybrid loopy belief propagation* In Proceedings of the Third European Workshop on Probabilistic Graphical Models (PGM'06), pages 317-324, Prague, Czech Republic, 2006

- Sun W. and Chang K. (2007) *Unscented Message Passing for Arbitrary Continuous Variables in Bayesian Networks*. Department of Systems Engineering and Operations Research, George Mason University, Fairfax.
- Yuan C. and Druzdzel M. J. (2007) *Importance Sampling for General Hybrid Bayesian Networks*. Department of Computer Science and Engineering, Mississippi State University, Mississippi, 2007.
- Koller, D. and Friedman N. (2009) *Probabilistic Graphical Models. Principles and Techniques*. The MIT Press, Cambridge, Massachusetts, 2009
- Sun, W., Chang K. and Laskey K. (2010) *Scalable Inference for Hybrid Bayesian Networks with Full Density Estimations*. The Sensor Fusion Lab, George Mason University, Fairfax, 2010
- Gogate V. and Detcher R. (2012) *Approximate Inference Algorithms for Hybrid Bayesian Networks with Discrete Constraints*. University of California, Irvine, 2012